

Locally Diverse Constellations from the Special Orthogonal Group

David A. Karpuk, *Member, IEEE*, Camilla Hollanti, *Member, IEEE*

Abstract

To optimize rotated, multidimensional constellations over a single-input, single-output Rayleigh fading channel, a family of rotation matrices is constructed for all dimensions which are a power of 2. This family is a one-parameter subgroup of the group of rotation matrices, and is located using a gradient descent scheme on this Lie group. The parameter defining the family is chosen to optimize the cutoff rate of the constellation. The optimal rotation parameter is computed explicitly for low signal-to-noise ratios.

These rotations outperform full-diversity algebraic rotations in terms of cutoff rate at low SNR (signal-to-noise ratio) and bit error rate at high SNR in dimension $n = 4$. However, a QAM (quadrature amplitude modulation) constellation rotated by such a matrix lacks full diversity, in contrast with the conventional wisdom that good signal sets exhibit full diversity. A new notion of diversity, referred to as local diversity, is introduced to attempt to account for this behavior. Roughly, a locally fully diverse constellation is fully diverse only in small neighborhoods. A local variant of the minimum product distance is also introduced and is shown experimentally to be a superior predictor of constellation performance than the minimum product distance in dimension $n = 4$.

Index Terms

Rayleigh fading channel, rotated constellations, non-uniform constellations, capacity, cutoff rate, Lie groups, full diversity

D. Karpuk and C. Hollanti are with the Department of Mathematics and Systems Analysis, Aalto University, Espoo, Finland. emails: {david.karpuk, camilla.hollanti}@aalto.fi

D. Karpuk is supported by Academy of Finland Postdoctoral grant #268364.

C. Hollanti is supported by Academy of Finland grants #276031, #282938, and #283262, and by a grant from Magnus Ehrnrooth Foundation, Finland. The support from the European Science Foundation under the COST Action IC1104 is also gratefully acknowledged.

I. INTRODUCTION

A. Rotated Constellations and the Cutoff Rate

Communication over a wireless channel typically involves transmitting a signal which is subsequently affected by fading and noise. Ensuring that such transmission is done reliably requires an energy-efficient constellation resistant to these effects. It is known that rotating constellations is an effective way to combat the effect of fading, while maintaining the energy of each codeword. Rotations effectively combat fading by increasing the diversity of the constellation.

Numerous algebraic and number-theoretic techniques exist to construct fully-diverse rotated cubic lattices with good minimum product distance [1], [2], [3], of which a finite subset is selected to serve as the constellation. These algebraic methods are effective at producing fully-diverse constellations which have good performance in terms of error rate. However, the rigidity of the algebraic constructions imposes a lack of flexibility in adjusting the particular rotation matrix in the presence of, for example, an outer error-correcting code or varying SNR (signal-to-noise ratio). Furthermore, the number-theoretic constructions of [3] assume high SNR and that the signal set is essentially infinite, that is, consists of an entire lattice in \mathbb{R}^n .

We investigate potential improvements in the algebraic method of [3] by dropping these assumptions and investigating rotations via a finer and SNR-dependent objective function controlling constellation performance. This objective function is an approximation of the capacity of the constellation known as the *cutoff rate*, which has a long history of being used as design criterion, but almost exclusively for erasure and AWGN channels; see the classical references [4], [5] and the references therein. It has been utilized much less for the fading channel model we consider, in favor of coarser, more algebraic measures of constellation performance.

Rotating constellations to improve CM (coded modulation) capacity has been investigated previously in [6], [7], [8]. Our work is partially inspired by [7], in which the authors approximate the CM capacity by the cutoff rate, and use the latter to construct good rotation matrices for 4-QAM and 16-QAM (quadrature amplitude modulation) constellations in \mathbb{C}^4 and \mathbb{C}^6 . The cutoff rate was also considered in conjunction with constellation design in [8], [9]. Our goal is to apply the general method of [7] of considering rotations which attempt to maximize the cutoff rate to the channel model of [3], for which one often constructs algebraic rotations. The cutoff rate for the channel we consider is a simple rational function of the coordinates of the constellation points

and therefore, unlike working with the CM capacity directly, does not require one to numerically compute numerically unstable unbounded integrals during optimization. In particular, the cutoff rate is simple enough to be susceptible to the methods of [10] for numerical optimization over rotation matrices, which we employ to investigate optimal rotation matrices.

B. Non-Uniform Constellations

By a *non-uniform* constellation, one typically means a “perturbation” of a standard QAM constellation, which heuristically behaves like a closer approximation of a Gaussian distribution than the standard “square” QAM constellations. Non-uniform constellations for the AWGN channel have been studied in [11], wherein the authors perform SNR-dependent optimization of such constellations, and in [12] with the goal of optimizing the BICM (bit interleaved coded modulation) capacity. Non-uniform constellations have been shown to have good error rate performance for the Gaussian [13] and Rayleigh [14] channels. Furthermore, non-uniform 64-QAM and 256-QAM constellations have been considered in DVB-NGH (Digital Video Broadcasting - Next Generation Handheld) implementation [15] due to the improvement in capacity, and are under consideration for inclusion in the ATSC (Advanced Television Systems Committee) 3.0 industry standard [16].

On the other hand, the DVB consortium has established rotated QAM constellations as a part of the DVB-T2 industry standard [17], and rotated QAM constellations in \mathbb{R}^2 and \mathbb{R}^4 as part of the DVB-NGH (next-generation handheld) standard [15]. A concise summary of such modulation techniques can be found in [18, §11]. A natural question is whether combining non-uniformity and rotations can offer an additional increase in capacity, and thus one would like to know how to find good rotations for arbitrary constellations, not just traditional M -QAM. As optimal non-uniform constellations can change with the SNR, and our goal is partly to rotate such constellations, we seek a rotation method which is SNR-dependent as well. Working with the CM capacity and cutoff rate also demands that one considers SNR-dependent rotations, since both of these objective functions depend on the SNR. As we shall see in Section III, optimal rotations can be highly SNR-dependent and optimizing rotations with respect to SNR can improve error rate performance.

C. Summary of Main Contributions

We present a general method for constructing good rotations of arbitrary constellations for single-input, single-output Rayleigh fast fading channels in dimensions which are a power of 2. Constructing rotations in arbitrarily high dimensions means abandoning explicit parameterizations of such matrices, since such parameterizations become non-canonical and unwieldy as the dimension of the ambient space increases. The collection of all rotations of \mathbb{R}^n is the *special orthogonal group* $SO(n)$ which has dimension $(n^2 - n)/2$ as a real manifold, meaning any parameterization of $n \times n$ rotation matrices requires at least $(n^2 - n)/2$ variables. We will present the necessary mathematical framework in Section V.

Our approach to constructing optimal rotation matrices differs largely from previous work on the subject, in that we effectively replace explicit parameterizations of rotation matrices with the matrix exponential map

$$\exp : \mathfrak{so}(n) \rightarrow SO(n) \quad (1)$$

where $\mathfrak{so}(n)$ is the Lie algebra of all skew-symmetric matrices. The general mathematical framework of Lie groups and Lie algebras allows us to construct well-performing families $\mathbf{Q}_{2^k}(t)$ of 2^k -dimensional rotation matrices for all k , which we view as one-parameter subgroups of $SO(2^k)$. The problem of optimizing over all $(n^2 - n)/2$ parameters defining a rotation matrix is thus reduced to optimizing over just a single parameter $t \in [0, \pi/2]$, which is easily done by exhaustive search and can be catered to the given SNR and constellation.

The flexibility of the parameter t is also beneficial if one wishes to further optimize due to the presence of an outer error-correcting code, a particular decoding algorithm, a particular bit interleaver, etc. In fact this is essentially what is done in the DVB-NGH standard [15, §6], which considers the (transpose of the) same family of matrices $\mathbf{Q}_4(t)$ which we construct explicitly, and optimizes the parameter t to suit the outer LDPC code rate and modulation order. By generalizing this construction to any dimension which is a power of 2, we hope to provide robust and adaptive families of rotation matrices to suit a number of applications. In general, we view our method as an alternative to the algebraic and number-theoretic constructions which accounts for the finiteness of the SNR, the boundedness of the constellation, and is adaptable to a number of potential objective functions which control system performance. We hope to generalize our method to the MIMO setting, and analogously improve upon the performance of

division algebra codes for MIMO Rayleigh fading channels [19], lattice codes for the Rician fading channel [20], and full-diversity unitary precoding techniques for MIMO channels [21]. Our adaptive approach may also be useful in newer protocols such as SCMA (sparse code multiple access) [22], in which algebraic rotations are proposed, but the complexity of SCMA codebook design may demand more intricate constructions in the future.

This article is intended to summarize, expand upon, and provide a more satisfying theoretical framework for the results of the conference papers [23], [24] by the current authors. In particular, the family of rotation matrices $\mathbf{Q}_{2^k}(t)$ we consider was discussed in [24] but we present here a simple, explicit formula in Section VI-B for their construction. We find in Section VI-C the optimal rotation within the family $\mathbf{Q}_{2^k}(t)$ in the low SNR regime and confirm our results via experiment. We investigate in Section IV-A the diversity order of our rotated constellations and show in Section VII-A that despite the lack of full diversity, they outperform the fully-diverse algebraically-rotated constellations of [3] in terms of both cutoff rate (at low SNR) and error performance (at high SNR) in dimension $2^k = 4$. We attribute this to a property we call *local full diversity*, which describes constellations which may not be fully diverse, but are so if one restricts to small neighborhoods. Our experiments indicate that constellations which enjoy this less stringent property can offer modest gains over fully diverse signal sets. This is in contrast with the conventional wisdom presented in [3] that optimal signal sets have full diversity.

II. CODED MODULATION CAPACITY AND CUTOFF RATE

We consider the general problem of constructing good constellations for Rayleigh fast fading single-input single-output (SISO) channels for transmission over n time instances. We assume perfect channel state information at the receiver, and as in [3, §2], by performing phase cancellation at the receiver, separating real and imaginary parts, and employing a bit interleaver, we can model the channel as

$$\mathbf{y} = \mathbf{h}\mathbf{x} + \mathbf{z} \quad (2)$$

where

- $\mathbf{x} = (x_1, \dots, x_n)^t \in \mathbb{R}^n$ is a column vector of transmitted data, selected uniformly at random from a finite constellation $\mathcal{X} \subset \mathbb{R}^n$ of size $|\mathcal{X}| = 2^q$,
- $\mathbf{h} = \text{diag}(h_1, \dots, h_n)$ is a real diagonal $n \times n$ matrix with h_i a Rayleigh random variable (i.e. the norm of a circularly symmetric complex Gaussian random variable) with $\mathbb{E}(h_i^2) = 1$,

- $\mathbf{z} = (z_1, \dots, z_n)^t \in \mathbb{R}^n$ is a noise vector with z_i a real zero-mean Gaussian random variable with variance N_0 , and
- $\mathbf{y} \in \mathbb{R}^n$ is the received vector.

We assume that the constellation is subject to an energy constraint of the form

$$\mathbb{E}||\mathbf{x}||^2 = \frac{1}{|\mathcal{X}|} \sum_{\mathbf{x} \in \mathcal{X}} ||\mathbf{x}||^2 = P \quad (3)$$

for some fixed P , where $|\mathcal{X}|$ denotes the cardinality of \mathcal{X} . Let us define the average energy per bit E_b by $P/\log_2(|\mathcal{X}|)$. We will present simulation results in terms of E_b/N_0 in decibels, but we will slightly abuse terminology and occasionally refer to this quantity as the SNR (signal-to-noise ratio) as well.

If \mathbf{T} is any $n \times n$ real matrix (in all cases we consider \mathbf{T} will either be a channel matrix or a rotation matrix), then by the notation $\mathbf{T}\mathcal{X}$ we mean the constellation

$$\mathbf{T}\mathcal{X} = \{\mathbf{T}\mathbf{x} \mid \mathbf{x} \in \mathcal{X}\} \subset \mathbb{R}^n. \quad (4)$$

A. Definitions and Basic Notions

In this section we recall some familiar formulas for the CM capacity of an n -dimensional constellation $\mathcal{X} \subset \mathbb{R}^n$ of size $|\mathcal{X}| = 2^q$. This discussion follows that of [7], but for the slightly different channel model (2). If we set $\mathbf{h} = \mathbf{I}_n$ in (2), the mutual information of the output \mathcal{Y} and the constellation \mathcal{X} is

$$I(\mathcal{Y}, \mathcal{X}) = \log_2(|\mathcal{X}|) - \sum_{\mathbf{x} \in \mathcal{X}} \int_{\mathbb{R}^n} \frac{\exp(-||\mathbf{y} - \mathbf{x}||^2/N_0)}{|\mathcal{X}|(\pi N_0)^{\frac{n}{2}}} \times \log_2 \left[\sum_{\substack{\mathbf{x}' \in \mathcal{X} \\ \mathbf{x}' \neq \mathbf{x}}} \exp\left(\frac{||\mathbf{y} - \mathbf{x}||^2 - ||\mathbf{y} - \mathbf{x}'||^2}{N_0}\right) \right] d\mathbf{y} \quad (5)$$

If $\mathbf{h} = \text{diag}(h_1, \dots, h_n)$ is a fixed fading matrix, then the conditional mutual information of \mathcal{Y} and \mathcal{X} given \mathbf{h} is

$$I(\mathcal{Y}, \mathcal{X}; \mathbf{h}) = I(\mathcal{Y}, \mathbf{h}\mathcal{X}). \quad (6)$$

One defines the *coded modulation capacity* (CM capacity) of \mathcal{X} by taking the expectation over all \mathbf{h} :

$$C^{\text{CM}}(\mathcal{X}) = \mathbb{E}_{\mathbf{h}} I(\mathcal{Y}, \mathcal{X}; \mathbf{h}) \quad (7)$$

For fixed n and E_b/N_0 , one would ultimately like to optimize $C^{\text{CM}}(\mathcal{X})$ over all constellations \mathcal{X} , but the large amount of numerical integration required to compute $C^{\text{CM}}(\mathcal{X})$ makes this intractable, especially with increasing dimension and constellation size. In fact, even evaluating $C^{\text{CM}}(\mathcal{X})$ with the numerical precision required to apply our optimization techniques seems out of reach, even for small constellations in small dimensions. Our first step is therefore to replace $C^{\text{CM}}(\mathcal{X})$ with a more tractable objective function.

B. The Cutoff Rate

It is known [7], [8], [25] that one can bound $I(\mathcal{Y}, \mathcal{X}; \mathbf{h})$ below by a quantity $R_0(\mathcal{X}; \mathbf{h})$, which for the model (2) is given by

$$R_0(\mathcal{X}; \mathbf{h}) = q - \log_2 \left[1 + 2^{-q} \sum_{\mathbf{x} \neq \mathbf{y} \in \mathcal{X}} \exp \left(\frac{-\|\mathbf{h}(\mathbf{x} - \mathbf{y})\|^2}{8N_0} \right) \right] \quad (8)$$

where recall that $|\mathcal{X}| = 2^q$, so that each constellation point encodes q bits. One can now define

$$R_0(\mathcal{X}) = \mathbf{E}_{\mathbf{h}} R_0(\mathcal{X}; \mathbf{h}) \quad (9)$$

and it follows trivially that $C^{\text{CM}}(\mathcal{X}) \geq R_0(\mathcal{X})$.

To establish numerical techniques which are independent of computing an expectation over all \mathbf{h} , one uses Jensen's inequality to further lower bound $R_0(\mathcal{X})$ by the *cutoff rate* $R(\mathcal{X})$, defined by

$$R(\mathcal{X}) = q - \log_2 \left[1 + 2^{-q} \sum_{\mathbf{x} \neq \mathbf{y} \in \mathcal{X}} \prod_{i=1}^n \frac{1}{1 + \frac{(x_i - y_i)^2}{8N_0}} \right] \quad (10)$$

For some fixed $\mathbf{x} \neq \mathbf{y} \in \mathcal{X}$, the corresponding summand in (10) is related to the pairwise error probability of confusing \mathbf{y} for \mathbf{x} at the receiver [2]. Our goal is now to optimize our constellation \mathcal{X} by treating $R(\mathcal{X})$ as an objective function we wish to optimize for a given E_b/N_0 , dimension n , and modulation order $|\mathcal{X}|$. We briefly remark that most authors [7], [25] use the term “cutoff rate” to refer to the quantity $R_0(\mathcal{X})$, but we reserve the term for $R(\mathcal{X})$ to simplify terminology, since it is the main objective function we consider.

Studying the cutoff rate eliminates the need to perform the numerical integration necessary to evaluate $C^{\text{CM}}(\mathcal{X})$, which can be computationally demanding or even infeasible, especially since the integral is unbounded and the integrand is numerically unstable. Furthermore, the use of Jensen's Inequality has eliminated computing $\mathbf{E}_{\mathbf{h}}(\cdot)$, which would further require computing a

second numerically unstable integral, or measuring $R_0(\mathcal{X})$ via Monte Carlo simulations. Even small improvements in $C^{\text{CM}}(\mathcal{X})$, $R_0(\mathcal{X})$, and $R(\mathcal{X})$ can introduce non-trivial benefits in terms of, for example, error performance, and small gains in the former two quantities are necessarily very difficult to measure, especially for larger dimensions and modulation orders. Hence we will almost exclusively measure performance in terms of $R(\mathcal{X})$. Note that the dependence of $R(\mathcal{X})$ on the noise variance N_0 allows us to tailor our constellations to particular SNR regimes.

III. NON-UNIFORM QAM CONSTELLATIONS IN \mathbb{R}^2

Let us begin by studying the model (2) when $n = 2$. The purpose of this section is to introduce non-uniform constellations and demonstrate their utility. Non-uniform QAM (NUQAM) constellations have been included in the DVB-NGH standard [15, §6] for modulation orders of $M = 64, 256$, wherein their defining parameters have been optimized according to the rate of an outer LDPC code. Here we demonstrate how optimal rotations of constellations vary when one changes both the SNR and the non-uniformity parameters, which helps motivate our more intricate constructions in later sections when we increase the value of n .

A. Optimal Non-Uniform Constellations

Let $q \geq 4$ and consider a set of parameters $\alpha = (\alpha_1, \dots, \alpha_{2^{q/2}-1}) \in \mathbb{R}^{2^{q/2}-1}$ with $\alpha_i > 0$ for all i . By an M -NUQAM constellation $\mathcal{X}(\alpha) \subset \mathbb{R}^2$ we mean the direct product of the set

$$\{-\alpha_{2^{q/2}-1}, \dots, -\alpha_1, \alpha_1, \dots, \alpha_{2^{q/2}-1}\} \quad (11)$$

with itself, such that $|\mathcal{X}(\alpha)| = M$. For example, when $q = 4$ and $\alpha = (1, 4)$, we have the 16-NUQAM constellation

$$\mathcal{X}(1, 4) = \{(\pm 1, \pm 1), (\pm 1, \pm 4), (\pm 4, \pm 1), (\pm 4, \pm 4)\}. \quad (12)$$

The α_i should be interpreted as perturbed versions of the normal parameters defining the standard M -QAM constellations. We will refer to the α_i as the *non-uniformity parameters*. Such constellations are known to increase capacity at lower SNR ranges [15].

For a fixed constellation size and E_b/N_0 , we optimize the non-uniformity parameters α by performing steepest descent on the objective function $R(\mathcal{X}(\alpha))$. In Table I we give examples of some such numerically-optimized NUQAM constellations, and in Fig. 1 we plot an example of a 1024-NUQAM constellation.

TABLE I
SOME EXAMPLE M -NUQAM CONSTELLATIONS

E_b/N_0 (dB)	M	optimal $\alpha = (\alpha_1, \dots, a_{2q/2-1})$ for M -NUQAM
8	16	(0.9732, 3.0088)
12	64	(0.9179, 2.7927, 4.8112, 7.2257)
15	256	(0.8912, 2.6844, 4.5119, 6.4022, 8.3956, 10.5573, 13.0147, 16.1037)

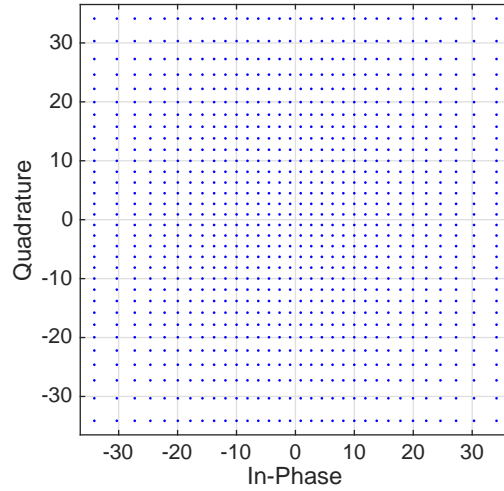


Fig. 1. Scatter plot of a 1024-NUQAM constellation in \mathbb{R}^2 , optimized for $E_b/N_0 = 20$ dB.

Of course, one could maximize the cutoff rate of \mathcal{X} by performing steepest descent on every coordinate of every signal point. However, when $|\mathcal{X}| = 2^{12}$, for example, this approach would require optimization over $2 \cdot 2^{12} = 8192$ real parameters, whereas optimizing with respect to non-uniformity parameters requires only $2^{12/2-1} = 32$ real parameters. In the next subsection we shall see how to further improve the cutoff rate by the addition of a single real parameter to the search space, namely one governing rotation.

B. Optimal Rotation Angles

Let us study how to further improve the cutoff rate of constellations in \mathbb{R}^2 by rotating them. While this offers large benefits in terms of rate, we note that rotated constellations are necessarily more difficult to decode, since one cannot decode the coordinates independently. We optimize

the rotation angle of a constellation \mathcal{X} by considering the objective function

$$R(\mathbf{Q}_2(t)\mathcal{X}), \quad \mathbf{Q}_2(t) = \begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix} \quad (13)$$

and performing a brute force search over the interval $t \in [0, \pi/2]$.

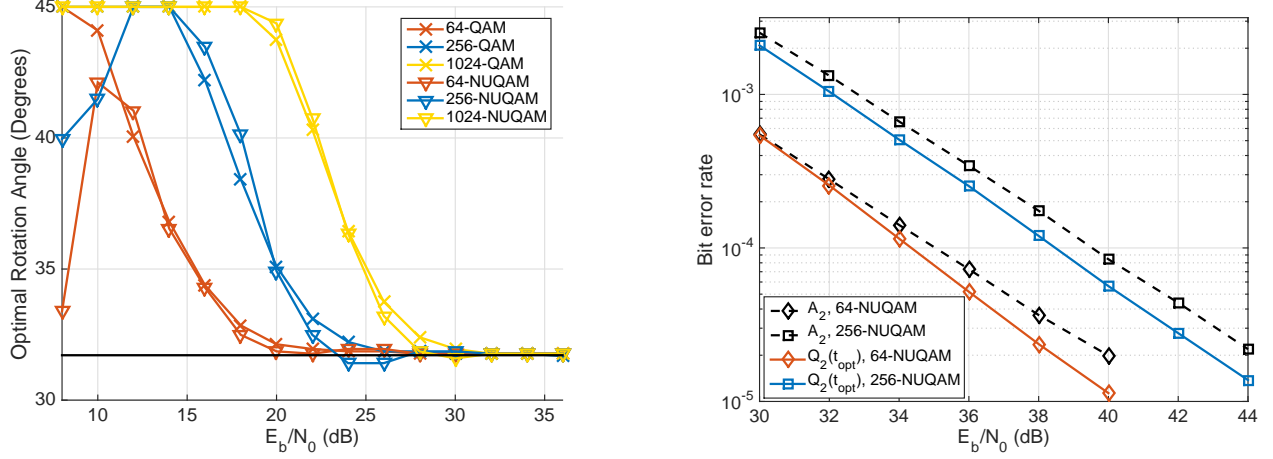


Fig. 2. On the left, we plot optimal rotation angles for M -QAM and M -NUQAM constellations in \mathbb{R}^2 , for $M = 64, 256$, and 1024. The horizontal black line represents the angle corresponding to the optimal two-dimensional algebraic rotation matrix \mathbf{A}_2 [26]. On the right, we plot the bit error rate of the 64- and 256-NUQAM constellations, rotated by the optimal rotation matrix $\mathbf{Q}_2(t_{\text{opt}})$ as well as the algebraic rotation \mathbf{A}_2 .

To motivate the need for constellation- and SNR-dependent rotations, we plot in Fig. 2 the rotation angle $t_{\text{opt}} \in [0, \pi/2]$ which optimizes the objective function in (13) as a function of E_b/N_0 . The NUQAM constellations used were optimized for each value of E_b/N_0 and then optimal rotation angles were considered. It is clear from Fig. 2 that the optimal rotation angle depends on the modulation order, the SNR, and the uniformity of the constellation.

Also in Fig. 2 we plot the bit error rate for the 64- and 256-NUQAM constellations, rotated by the optimal matrix $\mathbf{Q}_2(t_{\text{opt}})$ as well as the optimal algebraic rotation \mathbf{A}_2 [26]. Bit strings of length at least 10^8 were modulated to the constellation points at every value of E_b/N_0 , according to a Gray labeling. We see that the gain obtained by rotating on a per-SNR basis can be as high as 1.5 dB at high values of E_b/N_0 for both NUQAM constellations. While the operating range for the SNR is quite high, we are ultimately interested in comparison with the algebraic lattices of [19] which are designed for the very high SNR regime. Thus the main point is not the magnitude of

the gains but rather that there are any gains at all, especially in the very high regime for which algebraic rotations are often considered optimal.

These plots provide motivation for finding SNR- and constellation-dependent rotation matrices for constellations in \mathbb{R}^2 as well as in larger ambient spaces. The algebraic and number-theoretic methods of [3], [26] of constructing good rotated \mathbb{Z}^n -lattices necessarily fall somewhat short of this goal since NUQAM constellations are not subsets of lattices, and furthermore algebraic lattices are not SNR-dependent.

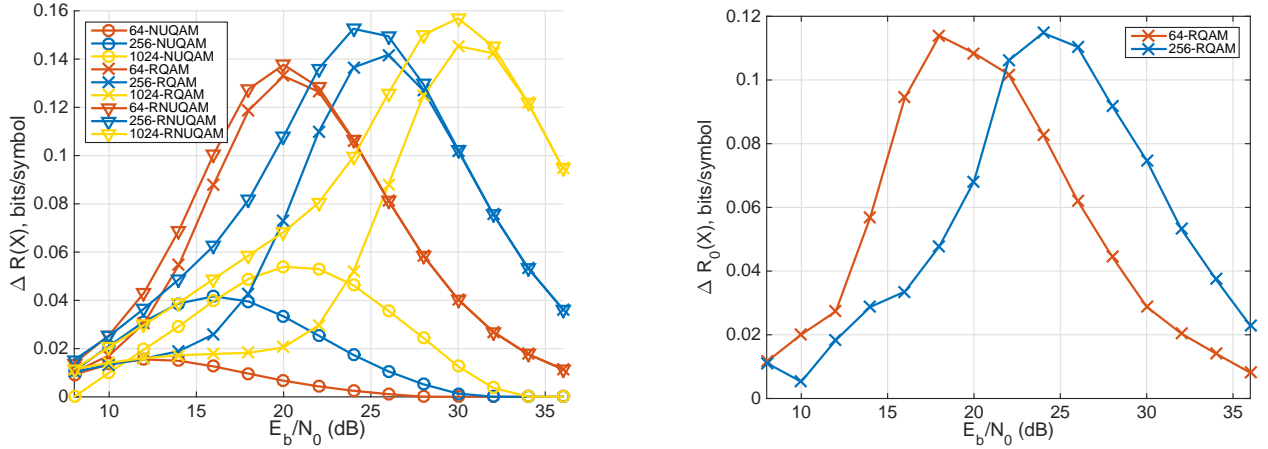


Fig. 3. On the left, we plot the improvement in $R(\mathcal{X})$ for non-uniform (NUQAM), rotated (RQAM), and rotated non-uniform (RNUQAM) constellations over their unrotated, uniform counterparts. On the right, we plot the experimentally measured increase in $R_0(\mathcal{X})$ for 64-RQAM and 256-RQAM constellations.

In Fig. 3 we plot the overall increases

$$\Delta R(\mathcal{X}) = R(\mathcal{X}) - R(\mathcal{X}_{\text{QAM}}), \quad \Delta R_0(\mathcal{X}) = R_0(\mathcal{X}) - R_0(\mathcal{X}_{\text{QAM}}) \quad (14)$$

in cutoff rate for various constellations \mathcal{X} , over their uniform, unrotated counterparts of the same modulation order, denoted by \mathcal{X}_{QAM} . In measuring the increase in $R(\mathcal{X})$, we used non-uniform (NUQAM), rotated (RQAM), and rotated non-uniform (RNUQAM) constellations, for $M = 64, 256$, and 1024 , over their uniform, unrotated counterparts. For the NUQAM and RQAM constellations, the non-uniformity and rotation parameters were numerically optimized by performing steepest descent on the cutoff rate. The RNUQAM constellations were first optimized with respect to the non-uniformity parameters, and then with respect to rotation. Notice that non-uniformity and rotation can improve the cutoff rate of the constellation at different values of

E_b/N_0 . For example at $E_b/N_0 = 20$ dB, 1024-QAM benefits much more from introducing non-uniformity than from rotation. Lastly we note that the benefit of rotation is most pronounced in the finite SNR regime.

The quantity $\Delta R_0(\mathcal{X})$ was computed for 64-RQAM and 256-RQAM by experimentally averaging the value of $R_0(\mathcal{X}; \mathbf{h}) - R_0(\mathcal{X}_{\text{QAM}}; \mathbf{h})$ over $2 * 10^5$ channels \mathbf{h} . Here optimal rotations were computed on a per-SNR basis, so that \mathcal{X} was chosen to be the rotated version of \mathcal{X}_{QAM} which maximizes $R(\mathcal{X})$. As one can see by comparing the plots, maximizing $R(\mathcal{X})$ results in a significant increase in $R_0(\mathcal{X})$ as well, and thus we include these simulation results to empirically demonstrate the strong correlation between these two objective functions. However, due to the difficulty of computing $R_0(\mathcal{X})$ even for low dimensions, we will forego measuring constellation performance in terms of $R_0(\mathcal{X})$ and only present further simulation results in terms of $R(\mathcal{X})$.

IV. LOCAL DIVERSITY

Note that in Fig. 2 the optimal angle for a number of constellations and SNR values was $t = 45^\circ$, which does *not* offer full diversity (in the sense of [3] and defined below) in the case of a rotated QAM constellation. Thus at finite SNR, constellations which are not fully diverse have the potential to outperform fully diverse constellations. In this section we define a subtler notion of diversity which seeks to explain this phenomenon.

A. Full Diversity and Minimum Product Distance

Let $\mathcal{X} \subset \mathbb{R}^n$ be any constellation. If we let $N_0 \rightarrow 0$, that is, consider the high SNR regime, then we could further approximate the cutoff rate (10) by

$$R(\mathcal{X}) \approx q - \log_2 [1 + 2^{-q} S_0] \quad (15)$$

where

$$S_0 = \sum_{\mathbf{x} \neq \mathbf{y} \in \mathcal{X}} \prod_{\substack{i=1 \\ x_i \neq y_i}}^n \frac{8N_0}{(x_i - y_i)^2} \quad (16)$$

To maximize $R(\mathcal{X})$, then, one wishes to minimize S_0 . As a first attempt at doing so, one maximizes the number of terms in each of the above products. For that we consider the *diversity order* L , defined by

$$L(\mathcal{X}) = \min_{\mathbf{x} \neq \mathbf{y} \in \mathcal{X}} \#\{i \mid x_i \neq y_i\} \quad (17)$$

In other words, $L(\mathcal{X})$ measures the minimum number of coordinates in which two signal points differ. When $L(\mathcal{X}) = n$ one says that \mathcal{X} has *full diversity*, which is often a desirable property and can be effectively achieved via algebraic methods [3].

As a sharper measure of constellation performance, one can then consider the *minimum product distance* $d_p(\mathcal{X})$ defined by

$$d_p(\mathcal{X}) = \min_{\mathbf{x} \neq \mathbf{y} \in \mathcal{X}} \prod_{\substack{i=1 \\ x_i \neq y_i}}^n |x_i - y_i| \quad (18)$$

At high SNR design criteria typically reduces to choosing an \mathcal{X} which maximizes $d_p(\mathcal{X})$, for which the number-theoretic constructions of [3] are effective.

B. Local Cutoff Rate and Local Diversity

The notion of full diversity can be somewhat restrictive. Indeed, it may not matter that much that all of the coordinates of \mathbf{x} and \mathbf{y} are distinct if they are very far from one another in Euclidean distance. Thus rate and error performance should depend only on local neighborhoods of the constellation and not on its global structure. We loosen the notion of full diversity in the following way to account for this fact.

Let $\mathbf{x} \in \mathcal{X}$ and fix a radius $0 < r \leq \infty$. Define the ball of radius r at \mathbf{x} by

$$B(\mathbf{x}, r) = \{\mathbf{y} \in \mathcal{X} \mid \|\mathbf{x} - \mathbf{y}\|_2 \leq r\}. \quad (19)$$

The intuition behind the following definition is that the summands which dominate the summation appearing in (10) are those for which the points $\mathbf{x} \neq \mathbf{y}$ are close in Euclidean distance.

Definition 1: Given a radius $0 < r \leq \infty$, we define the *local cutoff rate* $R(\mathcal{X}, r)$ of \mathcal{X} by

$$R(\mathcal{X}, r) = q - \log_2(1 + 2^{-q}S(r)) \quad (20)$$

where

$$S(r) = \sum_{\mathbf{x} \in \mathcal{X}} \sum_{\substack{\mathbf{y} \in B(\mathbf{x}, r) \\ \mathbf{y} \neq \mathbf{x}}} \prod_{i=1}^n \frac{1}{1 + \frac{(x_i - y_i)^2}{8N_0}}. \quad (21)$$

The accuracy of $R(\mathcal{X}, r)$ to predict the value of $R(\mathcal{X})$ depends on the value of r , and the r one wishes to consider may in turn depend on the SNR. For $r = \infty$ we have $R(\mathcal{X}, \infty) = R(\mathcal{X})$. One of the main advantages of considering local cutoff rates is that one can often find explicit provably

optimal rotations, as we shall see. These rotations will not result in fully diverse constellations, but they will be “locally” fully diverse in the following sense.

Definition 2: Given a radius $0 < r \leq \infty$, the *local diversity order* of \mathcal{X} is

$$L(\mathcal{X}, r) = \min_{\mathbf{x} \in \mathcal{X}} \min_{\substack{\mathbf{y} \in B(\mathbf{x}, r) \\ \mathbf{y} \neq \mathbf{x}}} \#\{i \mid x_i \neq y_i\}. \quad (22)$$

If $L(\mathcal{X}, r) = n$ for some radius r then we shall say that \mathcal{X} has *local full diversity*. If $L(\mathcal{X}, \infty) = n$ we shall say that \mathcal{X} has *global full diversity*.

The property of local full diversity of course depends on the particular radius, which we shall always make explicit. Global full diversity is equivalent to full diversity in the traditional sense of (17), but we append the adjective ‘global’ to emphasize the distinction.

For example, consider the constellation

$$\mathcal{X} = \{(\pm\sqrt{2}, 0), (0, \pm\sqrt{2})\} \subset \mathbb{R}^2. \quad (23)$$

One sees easily that $L(\mathcal{X}, 2) = 2$ but $L(\mathcal{X}, \infty) = 1$. Hence \mathcal{X} is locally fully diverse but not globally fully diverse. In fact, \mathcal{X} is the standard 4-QAM constellation in \mathbb{R}^2 rotated by $t = 45^\circ$, which the experiments of the previous section show is often an optimal rotation for constellations in \mathbb{R}^2 .

Loosely speaking, if the SNR is such that the large majority of errors cause constellation points to be confused only for others within a ball of radius r , then the performance of a locally fully diverse constellation should be indistinguishable from that of a globally diverse constellation. The above definitions reflect the fact that error performance is largely a local invariant of a constellation.

We can compare two locally diverse constellations at high SNR using the following variant of the minimum product distance.

Definition 3: Given a radius $0 < r \leq \infty$, the *local minimum product distance* of \mathcal{X} is

$$d_p(\mathcal{X}, r) = \min_{\mathbf{x} \in \mathcal{X}} \min_{\substack{\mathbf{y} \in B(\mathbf{x}, r) \\ \mathbf{y} \neq \mathbf{x}}} \prod_{\substack{i=1 \\ x_i \neq y_i}}^n |x_i - y_i| \quad (24)$$

For $r = \infty$ we have the obvious equality $d_p(\mathcal{X}, \infty) = d_p(\mathcal{X})$. In Section VII-A we will show that the local minimum product distance can sometimes predict the relative performance of two constellations when $d_p(\mathcal{X})$ cannot.

V. THE SPECIAL ORTHOGONAL GROUP $SO(n)$

The large gains in cutoff rate offered by rotations at finite SNR for two-dimensional constellations motivate us to find good rotations for higher-dimensional constellations, which are tailored for a given SNR and constellation. To provide the proper mathematical framework, we first discuss the natural habitat of higher-dimensional rotations.

A. Basic Definitions

To properly describe the families of rotation matrices we construct, we need to discuss the structure of the *special orthogonal group* $SO(n)$ of all n -dimensional rotation matrices. The group $SO(n)$ of rotations of \mathbb{R}^n is defined by

$$SO(n) = \{\mathbf{Q} \in GL(n) \mid \mathbf{Q}\mathbf{Q}^t = \mathbf{I}_n, \det(\mathbf{Q}) = 1\} \quad (25)$$

where $GL(n)$ is the group of all invertible real $n \times n$ matrices. The dimension of $SO(n)$ as a manifold is $(n^2 - n)/2$, which can be thought of as the minimum number of parameters required to describe an $n \times n$ rotation matrix. The special orthogonal group is a *Lie group*, which is both a group and a manifold such that the group operations are continuous with respect to the manifold structure. As a general reference for the theory of Lie groups we recommend [27].

The *Lie algebra* $\mathfrak{so}(n)$ of $SO(n)$ is the tangent space at the identity matrix $\mathbf{I}_n \in SO(n)$, and thus is a real vector space of dimension $(n^2 - n)/2$. We have the following convenient explicit description:

$$\mathfrak{so}(n) = \{\mathbf{A} \in M(n) \mid -\mathbf{A} = \mathbf{A}^t\} \quad (26)$$

where $M(n)$ is the set of all $n \times n$ real matrices. We pass from the Lie algebra to the Lie group using the exponential map $\exp : \mathfrak{so}(n) \rightarrow SO(n)$, defined by the familiar power series

$$\exp(\mathbf{A}) = \mathbf{I}_n + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \frac{\mathbf{A}^3}{3!} + \cdots \quad (27)$$

One can verify easily that the exponential series converges for all \mathbf{A} . There also exists a logarithm map $\log : SO(n) \rightarrow \mathfrak{so}(n)$, defined by a similar power series and satisfying $\exp(\log(\mathbf{A})) = \mathbf{A}$. It is a basic fact that the exponential of a skew-symmetric matrix is a rotation matrix, and that the logarithm of a rotation matrix is skew-symmetric.

A one-parameter subgroup of $SO(n)$ is the image of a continuous group homomorphism $\mathbb{R} \rightarrow SO(n)$. One can show all one-parameter subgroups are of the form

$$\mathbf{Q}(t) = \exp(t\mathbf{A}) \quad (28)$$

for some $\mathbf{A} \in \mathfrak{so}(n)$ (which depends on the particular subgroup) and $t \in \mathbb{R}$. The following lemma will prove useful for the particular one-parameter subgroups we will consider.

Lemma 1: Suppose that $\mathbf{A} \in M(n)$ is such that $\mathbf{A}^2 = -\mathbf{I}_n$, and let $\mathbf{Q}(t) = \exp(t\mathbf{A})$. Then

$$\mathbf{Q}(t) = \cos(t)\mathbf{I}_n + \sin(t)\mathbf{A}. \quad (29)$$

Proof: A straightforward computation using the definition of the exponential map yields

$$\exp(t\mathbf{A}) = \mathbf{I}_n + t\mathbf{A} + \frac{t^2\mathbf{A}^2}{2!} + \frac{t^3\mathbf{A}^3}{3!} + \cdots \quad (30)$$

$$= \left(1 - \frac{t^2}{2!} + \cdots\right) \mathbf{I}_n + \left(t - \frac{t^3}{3!} + \cdots\right) \mathbf{A} \quad (31)$$

$$= \cos(t)\mathbf{I}_n + \sin(t)\mathbf{A} \quad (32)$$

as claimed. ■

In the same way that $e^{ti} = \cos(t) + \sin(t)i$ where $i^2 = -1$ traces out the unit circle, the family $\exp(t\mathbf{A}) = \cos(t)\mathbf{I}_n + \sin(t)\mathbf{A}$ where $\mathbf{A} \in \mathfrak{so}(n)$, $\mathbf{A}^2 = -\mathbf{I}_n$ traces out a circle in the group $SO(n)$. Thus if one can find a desirable skew-symmetric matrix \mathbf{A} , optimizing over all of $SO(n)$ can be reduced to optimizing over a circle, which can be done by brute force.

While computing the exponential of an arbitrary matrix can be computationally demanding, the above proposition reduces us, in the case where $\mathbf{A}^2 = -\mathbf{I}_n$, to computing $\cos(t)$ and $\sin(t)$. This is useful when computing the optimal $t \in [0, 2\pi]$.

B. Gradient Fields and Steepest Descent on $SO(n)$

Let $f : M(n) \rightarrow \mathbb{R}$ be a differentiable real-valued function, and suppose that we prescribe coordinates \mathbf{q}_{ij} , $1 \leq i, j \leq n$ to the space $M(n)$. The gradient of f as a function on $M(n)$ can be identified with the vector field

$$\nabla f : M(n) \rightarrow M(n), \quad \nabla f(\mathbf{Q}) = \left(\frac{\partial f}{\partial q_{ij}}(\mathbf{Q}) \right)_{1 \leq i, j \leq n}, \quad \mathbf{Q} = (q_{ij})_{1 \leq i, j \leq n} \quad (33)$$

We wish to understand critical points of f when restricted to $SO(n)$, that is, describe the gradient field $\nabla(f|_{SO(n)})$. To that end, let us also define the vector field

$$X_f : SO(n) \rightarrow \mathfrak{so}(n), \quad X_f(\mathbf{Q}) = \nabla f(\mathbf{Q})\mathbf{Q}^t - \mathbf{Q}\nabla f(\mathbf{Q})^t \quad (34)$$

which is essentially a translated version of $\nabla(f|_{SO(n)})$, but is more convenient to work with since $X_f(\mathbf{Q})$ is skew-symmetric. The following theorem and its proof can be found in [10, §2.4], in which the authors work with a more general structure known as the Stiefel manifold. In the notation of [10], the Stiefel manifold is equal to the orthogonal group when $p = n$.

Theorem 1: [10, §2.4.4] Let $f : M(n) \rightarrow \mathbb{R}$ be a smooth, real-valued function. Then for any $\mathbf{Q} \in SO(n)$, we have

$$\nabla(f|_{SO(n)})(\mathbf{Q}) = X_f(\mathbf{Q})\mathbf{Q} \quad (35)$$

and thus \mathbf{Q} is a critical point of f if and only if $X_f(\mathbf{Q}) = 0$.

Proof: See [10]. ■

Let $f : M(n) \rightarrow \mathbb{R}$ be an objective function which we wish to minimize on the subset $SO(n) \subset M(n)$. To do this, we use the *conjugate gradient* method of [10, §3.4], also known as the method of *geodesic flow*. The conjugate gradient algorithm approximates the shortest path in $SO(n)$ between an initial point \mathbf{Q}_0 and an approximate local minimum \mathbf{Q}_N of f .

The conjugate gradient method presented in [10] works on a much more general class of manifolds, but we simplify and summarize the method briefly here for the specific case where the manifold is $SO(n)$. The conjugate gradient method starts with a rotation matrix \mathbf{Q}_0 and a step size h , and recursively updates via the formula

$$\mathbf{Q}_{k+1} = \exp(-hX_f(\mathbf{Q}_k)) \mathbf{Q}_k. \quad (36)$$

The matrix $X_f(\mathbf{Q}_k)$ is skew-symmetric and should be thought of as an infinitesimal gradient matrix living on the Lie algebra $\mathfrak{so}(n)$. Thus $\exp(-hX_f(\mathbf{Q}_k))$ is an element of $SO(n)$, which restricts this gradient search algorithm to the space of rotation matrices as desired. For an explanation of why this algorithm finds a local minimum of f , see [10]. Note that we do not assume f to be convex, and hence we must content ourselves with finding local rather than global minima of our objective functions.

Unlike many approaches to optimization over rotation matrices, the above numerical approach avoids the use of explicit parameterizations of rotation matrices. Such parameterizations become

unwieldy and non-canonical as the dimension n grows, hence we prefer the above method which circumvents any attempt at parameterizing the special orthogonal group as is done in [7].

VI. FAMILIES OF 2^k -DIMENSIONAL ROTATIONS

Given a fixed constellation $\mathcal{X} \subset \mathbb{R}^n$ and some E_b/N_0 , finding a rotated version of \mathcal{X} which optimizes the cutoff rate (10) is equivalent to finding maxima of the function $f : SO(n) \rightarrow \mathbb{R}$ defined by

$$f(\mathbf{Q}) = R(\mathbf{Q}\mathcal{X}) \quad (37)$$

In this section we will construct a family of candidates for good rotations for arbitrary constellations $\mathcal{X} \subset \mathbb{R}^{2^k}$ for any k . We first use the numerical conjugate gradient method of (36) to find example critical points, and then demonstrate how to explicitly construct such families of matrices in general, which approximate local maxima of the cutoff rate.

If $\mathcal{X} \subset \mathbb{R}^2$ is an M -QAM constellation or any rotated, non-uniform variant thereof, we will refer to $\mathcal{X}^n \subset \mathbb{R}^{2n}$ as a $2n$ D M -QAM constellation. Thus by a 4D 4-QAM constellation, for example, we mean the set

$$\mathcal{X} = \{(x_1, \dots, x_4) \in \mathbb{R}^4 \mid x_i = \pm 1\}. \quad (38)$$

of size 16. An M -QAM constellation in \mathbb{R}^{2n} thus contains M^n distinct signal points.

A. Finding Optimal Rotations Using Steepest Descent

Except for some trivial cases (for example $|\mathcal{X}| = 2$), analytically determining the critical points of f from (37) seems to be too difficult a task. On the other hand, the computational simplicity of evaluating the cutoff rate (as opposed to working directly with the CM capacity) allows one to use fairly complicated numerical schemes to find local maxima of f on $SO(n)$. Thus to obtain good rotations of non-uniform and high-dimensional constellations, we can use the conjugate gradient algorithm of the previous subsection.

Consider the 4D 4-QAM constellation of (38). For E_b/N_0 between 8 and 12 dB, we ran the conjugate gradient scheme (36) to numerically calculate optimal rotation matrices for this constellation. For $E_b/N_0 = 10$ dB, for example, the geodesic flow algorithm outputs a rotation

matrix \mathbf{Q}_N such that

$$\log(\mathbf{Q}_N) = \begin{bmatrix} 0 & 0.73 & 0.73 & 0.73 \\ -0.73 & 0 & 0.72 & -0.72 \\ -0.73 & -0.72 & 0 & 0.72 \\ -0.73 & 0.72 & -0.72 & 0 \end{bmatrix} \quad (39)$$

Here the initial point \mathbf{Q}_0 was chosen to be the matrix $\mathbf{Q}_0 = \exp(\mathbf{H})$, $\mathbf{H} = (h_{ij})$, $h_{ij} = 10^{-4}$ for $i > j$, $h_{ij} = -h_{ji}$ for $i < j$, and $h_{ii} = 0$ for all $i = 1, \dots, n$. Experiments suggest similar small perturbations of the identity matrix work well as initial points for the conjugate gradient method. Similarly structured matrices were obtained for other values of E_b/N_0 and other \mathcal{X} , namely, that $\log(\mathbf{Q}_N)$ appears to be a constant value times a matrix which depends only on the dimension n .

These numerically-obtained results suggest that to construct such matrices explicitly instead of running the gradient descent scheme (36) for every value of E_b/N_0 , one should construct a single skew-symmetric matrix \mathbf{A} and consider exponentiating all of its scalar multiples.

B. Explicit Construction of Rotation Matrices

We now demonstrate how to construct families of matrices of the form $\log(\mathbf{Q}_N)$ as in (39), when the ambient space has dimension a power of 2. To this end we will construct a one-parameter subgroup $\mathbf{Q}_{2^k}(t)$ of $SO(2^k)$, which in turn reduces the dimension of the search space from $\dim SO(n) = (n^2 - n)/2$ down to one. This saves significant computation time for dimensions as small as $n = 4$.

First let us recall the definition of the Hadamard matrices $\mathbf{H}_{2^k} \in M(2^k)$:

$$\mathbf{H}_1 = [1], \quad \mathbf{H}_{2^k} = \begin{bmatrix} \mathbf{H}_{2^{k-1}} & \mathbf{H}_{2^{k-1}} \\ \mathbf{H}_{2^{k-1}} & -\mathbf{H}_{2^{k-1}} \end{bmatrix} \quad (40)$$

We now construct skew-symmetric matrices $\mathbf{A}_{2^k} \in \mathfrak{so}(2^k)$ for $k \geq 1$ recursively in the following way:

$$\mathbf{B}_1 = [0], \quad \mathbf{B}_{2^k} = \begin{bmatrix} \mathbf{B}_{2^{k-1}} & \mathbf{H}_{2^{k-1}} \\ -\mathbf{H}_{2^{k-1}} & \mathbf{B}_{2^{k-1}} \end{bmatrix} \quad (41)$$

$$\mathbf{A}_{2^k} = (2^k - 1)^{-1/2} \mathbf{B}_{2^k} \quad (42)$$

The factor of $(2^k - 1)^{-1/2}$ is only a convenience that simplifies some expressions in what follows.

Given a fixed $n = 2^k$, we consider the one-parameter family of rotation matrices

$$\mathbf{Q}_n(t) = \exp(t\mathbf{A}_n) \in SO(n) \quad (43)$$

for $t \in \mathbb{R}$. The following proposition describes the matrices $\mathbf{Q}_n(t)$ explicitly, and relieves us of the computational burden of computing a matrix exponential.

Proposition 1: For any dimension $n = 2^k$ we have

$$\mathbf{Q}_n(t) = \cos(t)\mathbf{I}_n + \sin(t)\mathbf{A}_n. \quad (44)$$

Proof: One can show easily by induction that the matrices \mathbf{B}_n and \mathbf{H}_n anti-commute, that is, that

$$\mathbf{B}_n\mathbf{H}_n = -\mathbf{H}_n\mathbf{B}_n \quad (45)$$

for all $n \geq 2$. Furthermore, it is well-known and easy to show that the Hadamard matrices satisfy $\mathbf{H}_n^2 = n\mathbf{I}_n$. It is now straightforward to show that $\mathbf{B}_n^2 = -(n-1)\mathbf{I}_n$, or in other words that

$$\mathbf{A}_n^2 = -\mathbf{I}_n. \quad (46)$$

The result now follows from Lemma 1. ■

The previous proposition also allows us to restrict the parameter space to $t \in [0, \pi/2]$, as every matrix $\mathbf{Q}_n(t)$ for $t \in (\pi/2, 2\pi]$ can be obtained from a matrix $\mathbf{Q}_n(t)$ for $t \in [0, \pi/2]$ via negation and taking transposes, operations which do not affect constellation performance. For any constellation $\mathcal{X} \subset \mathbb{R}^n$ and a fixed value of E_b/N_0 , we can now optimize the cutoff rate with respect to rotation by computing

$$t_{\text{opt}} = \operatorname{argmax}_{t \in [0, \pi/2]} R(\mathbf{Q}_n(t)\mathcal{X}) \quad (47)$$

by simple exhaustive search.

When $n = 2$ we have

$$\mathbf{Q}_2(t) = \begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix} \quad (48)$$

and hence the family $\mathbf{Q}_2(t)$ is simply the whole group $SO(2)$. Finding the optimal rotation simply reduces to exhaustive search over the interval $t \in [0, \pi/2]$, which was exactly the procedure carried out in Section III-B.

When $n = 4$ one can see easily that $\mathbf{Q}_4(0.72)$ is approximately the matrix \mathbf{Q}_N from (39). The transpose of the matrix $\mathbf{Q}_4(0.56)$ was considered in conjunction with \mathcal{X} a 4D 4-QAM

constellation in the DVB-NGH standard [15] to minimize the bit error rate at the demapper. In a sense, one could describe our method as a generalization of the rotation matrices considered by DVB-NGH.

C. Optimal Rotations at Low SNR

Let us consider an arbitrary M -QAM constellation $\mathcal{X} \subset \mathbb{R}^n$ for $n = 2^k$. Our goal in this subsection is to find good rotations among the family $\mathbf{Q}_n(t)$ when the SNR is low. At low SNR, errors are mostly caused by large noise vectors as opposed to deep fades. One make this more precise by choosing $\mathbf{x} \neq \mathbf{y} \in \mathcal{X}$ and expanding the reciprocal of the product in the expression (10) for the cutoff rate. We obtain

$$\prod_{i=1}^n \left(1 + \frac{(x_i - y_i)^2}{8N_0} \right) = 1 + \frac{\|\mathbf{x} - \mathbf{y}\|_2^2}{8N_0} + \mathcal{O} \left(\frac{1}{N_0^2} \right) \quad (49)$$

Thus for large N_0 , the dominant term in the above is proportional to $\|\mathbf{x} - \mathbf{y}\|_2^2$, and the dominant terms in the sum (10) are those for which the Euclidean distance is smallest. Hence if we transmit \mathbf{x} and decode $\hat{\mathbf{x}} \neq \mathbf{x}$, it is overwhelmingly likely that $\|\mathbf{x} - \hat{\mathbf{x}}\|_2 = 2$, which is the minimal Euclidean distance between any two QAM constellation points.

The previous paragraph suggests that it suffices to consider the approximation

$$R(\mathcal{X}) \approx R(\mathcal{X}, 2) \quad (50)$$

of the cutoff rate by the local cutoff rate. Moreover, this approximation is convenient because it allows us to compute the rotation parameter t optimizing $R(\mathbf{Q}_n(t)\mathcal{X}, 2)$ explicitly.

Proposition 2: Let $\mathcal{X} \subset \mathbb{R}^n$ be a QAM constellation and let $n = 2^k$. The function $f : [0, \pi/2] \rightarrow \mathbb{R}$ defined by

$$f(t) = R(\mathbf{Q}_n(t)\mathcal{X}, 2) \quad (51)$$

has a single maximum at $t = \arccos(1/\sqrt{n})$.

Proof: Let $\mathbf{Q} \in SO(n)$ be any rotation matrix and consider the summation inside $R(\mathbf{Q}\mathcal{X}, r)$ defined by

$$S(r) = \sum_{\mathbf{x} \in \mathbf{Q}\mathcal{X}} \sum_{\substack{\mathbf{y} \in B(\mathbf{x}, r) \\ \mathbf{y} \neq \mathbf{x}}} \prod_{i=1}^n \frac{1}{1 + \frac{(x_i - y_i)^2}{8N_0}}. \quad (52)$$

For a fixed $\mathbf{x} \in \mathcal{X}$, we have $\mathbf{y} \in B(\mathbf{x}, 2)$ if and only if $\mathbf{x} - \mathbf{y} = \pm 2\mathbf{e}$ for some standard basis vector $\mathbf{e} \in \mathbb{R}^n$. Note that if $\mathbf{e} = \mathbf{e}_j$ where $j = 1, \dots, n$ then $\mathbf{Q}\mathbf{x} - \mathbf{Q}\mathbf{y}$ is the j^{th} column of \mathbf{Q} ,

whose i^{th} coordinate is q_{ij} . Since \mathbf{Q} is orthogonal and hence distance-preserving, it follows that $\mathbf{Q}B(\mathbf{x}, r) = B(\mathbf{Q}\mathbf{x}, r)$ and hence

$$S(r) = \sum_{\mathbf{x} \in \mathcal{X}} \sum_{\substack{\mathbf{y} \in B(\mathbf{x}, r) \\ \mathbf{y} \neq \mathbf{x}}} \prod_{i=1}^n \frac{1}{1 + \frac{q_{ij}^2}{2N_0}} \quad (53)$$

By the symmetry of the constellation, one can see that for all j the product $\prod_{i=1}^n \frac{1}{1 + \frac{q_{ij}^2}{2N_0}}$ appears in the above summand the same number of times. Now setting $\mathbf{Q} = \mathbf{Q}_n(t)$ we obtain for some positive integer c the expression

$$S(r) = c \sum_{j=1}^n \prod_{i=1}^n \frac{1}{1 + \frac{q_{ij}^2}{2N_0}} = \frac{cn}{g(t)} \quad (54)$$

where $g(t)$ is the function

$$g(t) = \left(1 + \frac{\cos^2(t)}{2N_0}\right) \left(1 + \frac{\sin^2(t)}{(n-1)2N_0}\right)^{n-1}. \quad (55)$$

We see that maximizing $R(\mathbf{Q}_n(t)\mathcal{X}, 2)$ with respect to t is equivalent to maximizing $g(t)$. One can compute the derivative of $g(t)$ explicitly to see that it has a single maximum on the interval $t \in [0, \pi/2]$ at $t = \arccos(1/\sqrt{n})$ which completes the proof. \blacksquare

We should expect from Proposition 2 that $t = \arccos(1/\sqrt{n})$ is a decent approximation of the optimal t at low SNR, whenever $R(\mathcal{X}, 2)$ is a decent approximation of $R(\mathcal{X})$. This is confirmed via experiment for $n = 2$ in Fig. 2 and as we shall see in the next section, for $n = 4$ in Fig. 4 and for $n = 8$ in Fig. 7.

D. Diversity of the Constellations $\mathbf{Q}_n(t)\mathcal{X}$

Here we present a short proof of the fact that if $\mathcal{X} \subset \mathbb{R}^n$ is a QAM constellation, then the rotated constellation $\mathbf{Q}_n(t)\mathcal{X}$ is locally fully diverse for $r = 2$ and all $t \in (0, \pi/2)$. Furthermore, we show by example that these rotated constellations lack full diversity. We will show in the next section that our locally fully diverse constellations can outperform the globally diverse constellations defined by number-theoretic rotations of [3], [26], even though our rotated constellations are not globally fully diverse.

Proposition 3: Let $\mathcal{X} \subset \mathbb{R}^n$ be a QAM constellation and let $\mathbf{Q} = (q_{ij})_{1 \leq i, j \leq n} \in SO(n)$ be any rotation matrix. Then $\mathbf{Q}\mathcal{X}$ satisfies $L(\mathbf{Q}\mathcal{X}, 2) = n$ if and only if $q_{ij} \neq 0$ for all i, j .

Proof: If $\mathbf{x} \in \mathcal{X}$ and $\mathbf{y} \in B(\mathbf{x}, 2)$ is not equal to \mathbf{x} , then $\mathbf{x} - \mathbf{y} = 2\mathbf{e}$ for some standard basis vector \mathbf{e} . Conversely, every standard basis vector of \mathbb{R}^n can be expressed as $\mathbf{e} = \frac{1}{2}(\mathbf{x} - \mathbf{y})$

for some $\mathbf{x} \in \mathcal{X}$ and $\mathbf{y} \in B(x, 2)$. Hence $\mathbf{Q}\mathcal{X}$ being locally fully diverse for radius $r = 2$ is equivalent to no entry of $\mathbf{Q}\mathbf{e}$ being zero for any standard basis vector \mathbf{e} . But this is exactly the statement that $q_{ij} \neq 0$ for all i, j . \blacksquare

For any choice of $t \in (0, \pi/2)$, it is now clear from Propositions 1 and 3 that for $\mathcal{X} \subset \mathbb{R}^n$ a QAM constellation with $n = 2^k$, the rotated version will satisfy

$$L(\mathbf{Q}_n(t)\mathcal{X}, 2) = n \quad (56)$$

that is, the rotated constellation will be locally fully diverse. In fact, Proposition 3 essentially says that it is very easy to achieve local full diversity; simply pick a rotation matrix without any zero entries, and such matrices are dense in $SO(n)$. Thus the property of local full diversity is not particular to the family $\mathbf{Q}_n(t)$; rather, among all rotations which offer local full diversity, we study these because they were so frequently the outcome of the numerical algorithm (36).

However, the rotations $\mathbf{Q}_n(t)$ do *not* offer full diversity when $n > 2$. Let us take, for example, $n = 4$, and consider $\mathbf{z} = \mathbf{x} - \mathbf{y}$ where $\mathbf{x} \neq \mathbf{y} \in \mathcal{X}$. Then

$$\mathbf{Q}_n(t) \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} az_1 + b(z_2 + z_3 + z_4) \\ az_2 + b(-z_1 + z_3 - z_4) \\ az_3 + b(-z_1 - z_2 + z_4) \\ az_4 + b(-z_1 + z_2 - z_3) \end{bmatrix} \quad (57)$$

where $a = \cos(t)$ and $b = \sin(t)/\sqrt{3}$. Now notice that for $z_1 = z_2 = 0$ and $z_3 = -z_4$ we have $(\mathbf{Q}_n(t)\mathbf{z})_1 = 0$. It is easy to check that the points $\mathbf{x} = (1, 1, 1, -1)^t$ and $\mathbf{y} = (1, 1, -1, 1)^t$ belonging to the standard 4D 4-QAM constellation will satisfy $(\mathbf{Q}_n(t)\mathbf{z})_1 = 0$, resulting in a rotated constellation that is not fully diverse.

VII. SIMULATION RESULTS FOR $n = 4$ AND $n = 8$

In this section we confirm the benefits of the family $\mathbf{Q}_n(t)$ of rotation matrices for $n = 4, 8$. Our basis for comparison is always the optimal algebraic rotation of the same dimension, as found in [26]. With regard to an algebraic rotation, “optimal” means with respect to the normalized minimum product distance of the rotated constellation. Maximum-likelihood decoding was used for all error rate simulations.

A. Rotated Constellations in \mathbb{R}^4

Let us further examine the family (43) for the simplest non-trivial example, that of rotated constellations in \mathbb{R}^4 . It follows from Proposition 1 that

$$\mathbf{Q}_4(t) = \cos(t)\mathbf{I}_4 + \frac{\sin(t)}{\sqrt{3}} \begin{bmatrix} 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & -1 \\ -1 & -1 & 0 & 1 \\ -1 & 1 & -1 & 0 \end{bmatrix} \quad (58)$$

We note that this one-parameter subgroup of $SO(n)$ was implicitly considered by the DVB-NGH standard [15, §6.4] when rotating 4D 4-QAM constellations. The DVB consortium specifies the particular rotation parameter t by instead specifying the parameter $\sin(t)$, but this difference is immaterial.

For 4D M -QAM and M -NUQAM and $M = 4, 16, 64$, we computed the optimal rotation matrix $\mathbf{Q}_4(t_{\text{opt}})$ by exhaustive search over the interval $t \in [0, \pi/2]$ as in (47) and plotted t_{opt} in degrees as a function of E_b/N_0 in Fig. 4. When $M = 64$, note the range of relatively low E_b/N_0 values for which $t_{\text{opt}} = \arccos(1/\sqrt{4}) = 60^\circ$, as predicted by Proposition 2. The NUQAM constellations in this simulation were first optimized on a per-SNR basis, and then optimal rotations were computed.

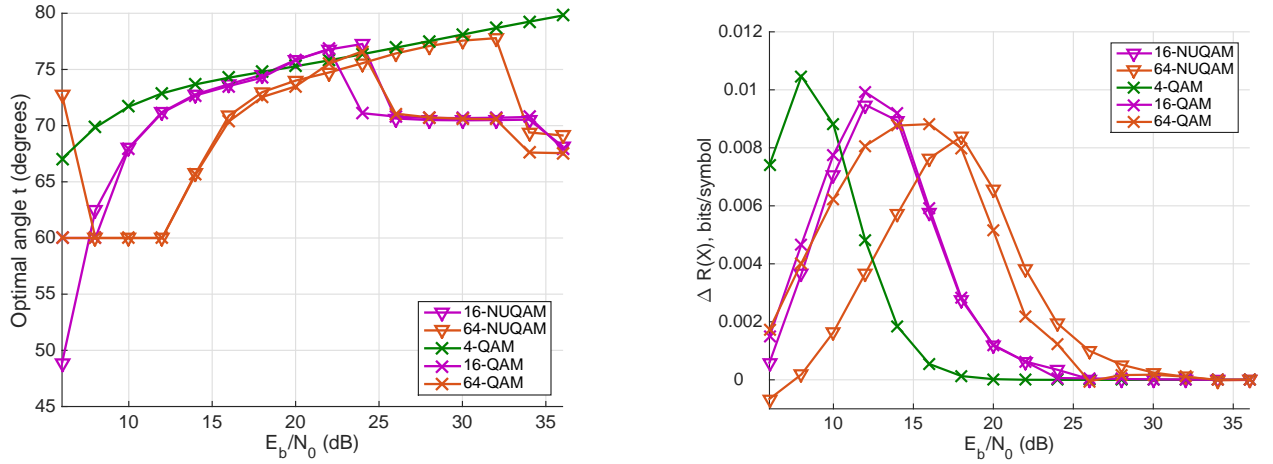


Fig. 4. On the left, the optimal rotation parameter t_{opt} as in (47) as a function of E_b/N_0 , for 4D M -QAM and M -NUQAM, $M = 4, 16, 64$. On the right, the improvement in $R(X)$ for 4D M -QAM and M -NUQAM constellations for $M = 4, 16$, and 64 when using the rotation matrices $\mathbf{Q}_4(t_{\text{opt}})$ as opposed to the Krüskemper rotation, which is the optimal algebraic rotation in \mathbb{R}^4 [26].

To compare our method with the optimal number-theoretic construction, we also plot in Fig. 4 the improvement

$$\Delta R(\mathcal{X}) = R(\mathbf{Q}_4(t_{\text{opt}})\mathcal{X}) - R(\mathbf{K}_4\mathcal{X}) \quad (59)$$

of our method over the algebraic method. Here \mathbf{K}_4 is the Krüskemper rotation, which is the optimal algebraic rotation of \mathbb{R}^4 and whose generator matrix is available at [26]. Our rotations $\mathbf{Q}_4(t_{\text{opt}})$ offer modest but consistent improvement in cutoff rate over the Krüskemper rotation at almost every value of E_b/N_0 and every $\mathcal{X} \subset \mathbb{R}^4$ we investigated. Despite the small improvement in cutoff rate at high SNR, this is the regime where the rotations $\mathbf{Q}_4(t_{\text{opt}})$ offer the highest error performance gains over \mathbf{K}_4 , as we see below.

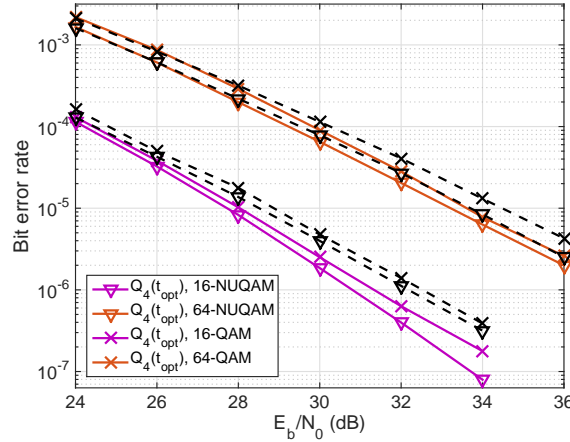


Fig. 5. Bit error rates for the family $\mathbf{Q}_4(t_{\text{opt}})$ for M -QAM and M -NUQAM in \mathbb{R}^4 , for $M = 16, 64$. The Krüskemper rotation, which is the optimal algebraic rotation in \mathbb{R}^4 , served as the basis for comparison and is plotted with the analogous dashed black curves. Simulations were performed by modulating approximately 10^9 bits at each value of E_b/N_0 .

We plot in Fig. 5 bit error rates of the family $\mathbf{Q}_4(t_{\text{opt}})$ for M -QAM and M -NUQAM, for $M = 16, 64$. Bit strings were assigned according to the Gray labeling of the underlying QAM symbols. The rotations $\mathbf{Q}_4(t_{\text{opt}})$ outperform the Krüskemper rotation for both uniform and non-uniform constellations at all values of E_b/N_0 . When $M = 16$, we see an increase in performance of up to 2 dB for non-uniform constellations and about 1.5 dB for uniform constellations. When $M = 64$, the increase is approximately 1 dB for uniform constellations, with a smaller increase in performance for non-uniform constellations.

While the operating SNR is very high, we have provided plots in this regime to study asymptotic behavior (i.e. diversity). This allows for fair comparison with the algebraic rotations

of [19] which are specifically tailored for the asymptotic regime. The crucial point is not the magnitude of the gains over the fully-diverse algebraic rotations, but rather that there are any gains at all in the absence of full diversity even in the asymptotic regime. Thus Fig. 5 underscores the original theoretical observation motivating local diversity, namely that full diversity is too restrictive a notion even in the very high SNR regime. In fact, one can show easily that $L(\mathbf{Q}_4(t_{\text{opt}})\mathcal{X}, \infty) = 3$ for all t_{opt} in the given E_b/N_0 range, that is that the rotations outperform the Krüskemper rotation even though their global diversity is in some sense deficient.

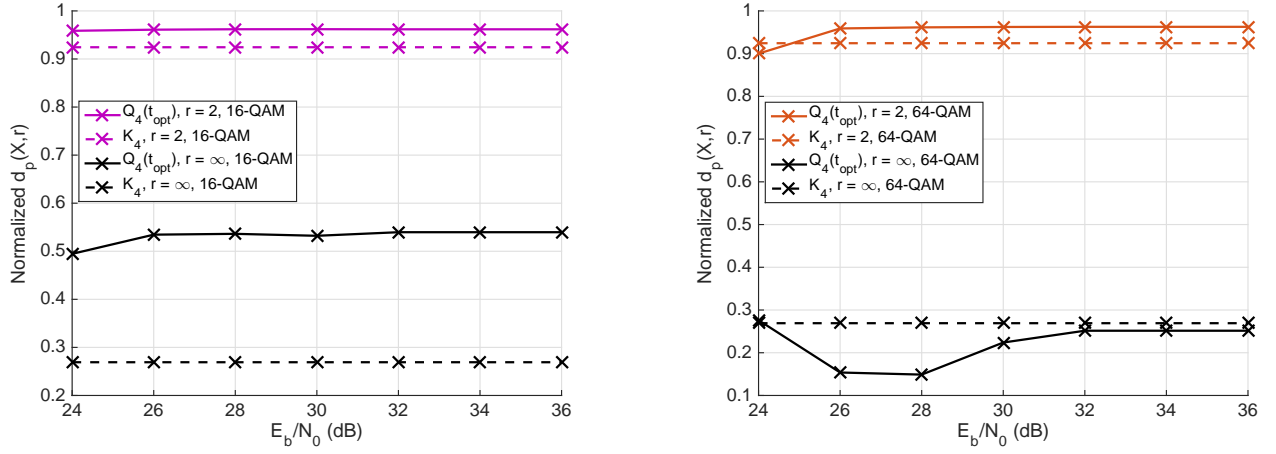


Fig. 6. Local (with $r = 2$) and global minimum product distances $d_p(\mathcal{X}, r)$ for the constellations rotated algebraically with the matrix \mathbf{K}_4 , and with the family $\mathbf{Q}_4(t_{\text{opt}})$. In the left figure are the results for 4D 16-QAM, and in the right figure are the results for 4D 64-QAM.

Finally in Fig. 6 we plot the various local and global minimum product distances $d_p(\mathbf{Q}\mathcal{X}, r)^{1/4}$ for $r = 2, \infty$ as defined in Section IV-B, for both $\mathbf{Q} = \mathbf{K}_4$ and $\mathbf{Q} = \mathbf{Q}_4(t_{\text{opt}})$, and $\mathcal{X} = 4\text{D } M\text{-QAM}$ for $M = 4, 16$. We see that for $M = 16$, both minimum product distances correctly predict that $\mathbf{Q}_4(t_{\text{opt}})$ outperforms \mathbf{K}_4 . However when $M = 64$, the local minimum product distance correctly predicts that the family $\mathbf{Q}_4(t_{\text{opt}})$ outperforms the rotation \mathbf{K}_4 for almost all values of E_b/N_0 , whereas the global minimum product distance effectively makes a mistake for almost all values of E_b/N_0 . Hence the local minimum product distance seems a more appropriate measure of the performance of these constellations.

B. Rotated Constellations in \mathbb{R}^8

In this subsection we study rotations of the 8D 4-QAM constellation $\mathcal{X} \subset \mathbb{R}^8$. In Fig. 7 we plot the value of t_{opt} as a function of E_b/N_0 . As predicted by Proposition 2, we obtain $t_{\text{opt}} = \arccos(1/\sqrt{8}) = 69.2952^\circ$ for several low values of E_b/N_0 , approximately in the range 4-7 dB. We also plot in Fig. 7 the improvement

$$\Delta R(\mathcal{X}) = R(\mathbf{Q}_8(t_{\text{opt}})\mathcal{X}) - R(\mathbf{C}_8\mathcal{X}) \quad (60)$$

where \mathbf{C}_8 is the optimal algebraic rotation in \mathbb{R}^8 [26]. We see from the plot that in the range of $E_b/N_0 = 4\text{-}14$ dB, the family $\mathbf{Q}_8(t_{\text{opt}})$ shows modest but consistent improvement over \mathbf{C}_8 in terms of cutoff rate, while the algebraic rotation has better cutoff rate outside of this region.

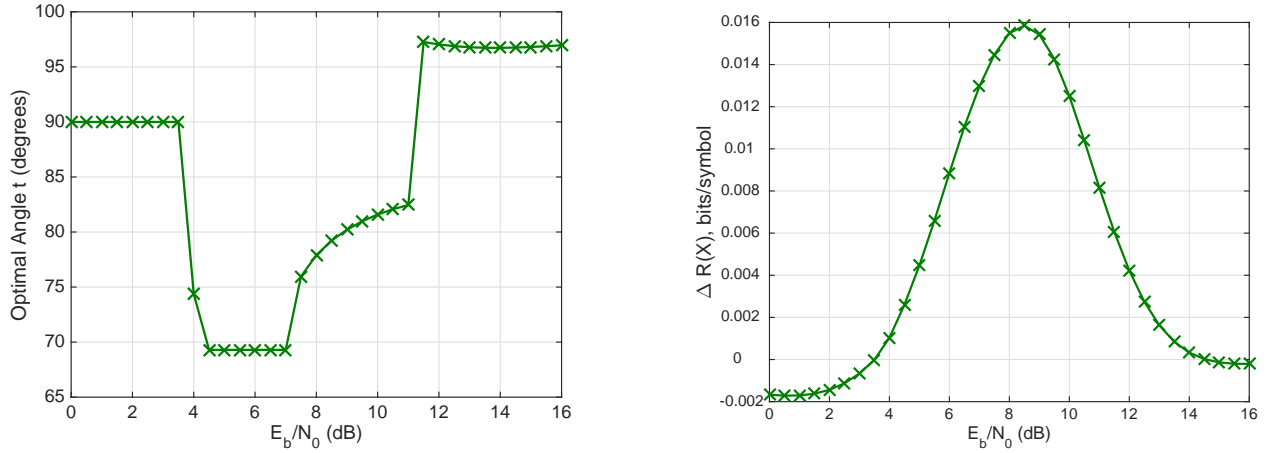


Fig. 7. On the left, optimal rotation parameter t_{opt} as in (47) as a function of E_b/N_0 , for 8D 4-QAM. On the right, improvement in $R(\mathcal{X})$ for 8D 4-QAM when using $\mathbf{Q}_8(t_{\text{opt}})$ as opposed to the algebraic rotation \mathbf{C}_8 of [26], which is the best-known algebraic rotation of \mathbb{R}^8 .

In Fig. 8 we plot the bit error rate of the rotations $\mathbf{Q}_8(t_{\text{opt}})$ as well as those of the optimal algebraic rotations of \mathbb{R}^n for $n = 5, 6, 7, 8$, denoted by $\mathbf{C}_5, \mathbf{M}_6, \mathbf{K}_7, \mathbf{C}_8$, respectively [26]. In the low SNR regime the performance of the two schemes is identical. Beyond $E_b/N_0 > 12$ dB (corresponding to a bit error rate of approximately 10^{-3}) the algebraic rotation outperforms the rotations $\mathbf{Q}_8(t_{\text{opt}})$.

While the rotations $\mathbf{Q}_8(t_{\text{opt}})$ have inferior error rate performance compared to \mathbf{C}_8 at high SNR, it can be shown that the global diversity of our scheme is only $L(\mathbf{Q}_8(t_{\text{opt}})\mathcal{X}, \infty) = 5$ whereas the fully-diverse algebraic rotations of \mathbb{R}^n achieve global diversity n . Thus the error plots show

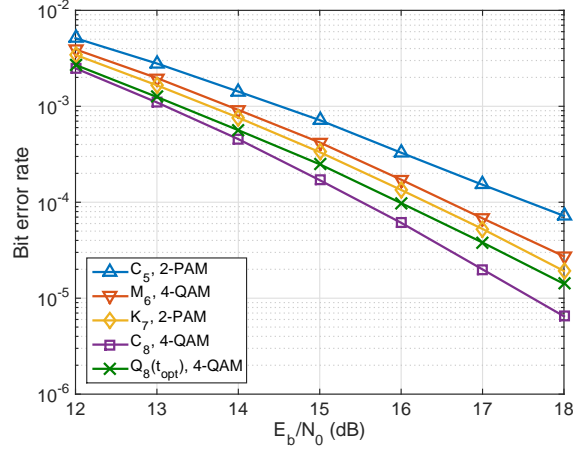


Fig. 8. Bit error rates for the family $\mathbf{Q}_8(t_{\text{opt}})$ for 4-QAM in \mathbb{R}^8 . The optimal algebraic rotations of \mathbb{R}^n for $n = 5, 6, 7, 8$, denoted respectively by $\mathbf{C}_5, \mathbf{M}_6, \mathbf{K}_7, \mathbf{C}_8$, served as the basis for comparison [26]. Simulations were performed by modulating approximately 10^8 bits at each value of E_b/N_0 according to the Gray labeling on the underlying per-coordinate constellation. A 2-PAM (pulse amplitude modulation) constellation was used per real axis for the odd-dimensional constellations, to ensure fair comparison across dimensions.

that $\mathbf{Q}_8(t_{\text{opt}})$ outperforms rotations whose global diversity is up to $n = 7$, lending credence to the original theoretical notion that local diversity should be a superior predictor of performance.

Lastly, we mention that neither local nor global minimum product distances appeared to predict the relative error performance, hence we omit plots of these quantities.

VIII. CONCLUSIONS AND FUTURE WORK

We have constructed a family of rotation matrices $\mathbf{Q}_{2^k}(t) \in SO(2^k)$ for every k with the goal of maximizing the cutoff rate of arbitrary constellations in \mathbb{R}^n . Our approach is an adaptive per-SNR optimization, in which an optimal rotation for a given constellation and SNR value can be computed by a simple exhaustive search on the interval $[0, \pi/2]$. Our rotations are applicable to non-uniform and uniform constellations alike, thereby improving the ability of non-uniformity to improve the cutoff rate. The cutoff rate has essentially served as a proxy for, and lower bounds, the CM capacity of the constellation. The computational intractability of the CM capacity for the channel under consideration makes the cutoff rate a much more attractive objective function.

Our rotations produce constellations which, while lacking full diversity, are “locally” fully diverse in the sense that every point is distinguishable from all of its closest neighbors by a

single coordinate. We have thus defined a new notion of diversity, deemed *local diversity*, which apparently suffices to guarantee modest improvement in both cutoff rate and error performance over the number-theoretic methods of [3] when the dimension is $n = 4$. By constructing a well-performing one-parameter subgroup of rotation matrices, we have added a degree of flexibility to the use of rotated constellations. We have used this parameter to optimize with respect to SNR, but one may use it to optimize with respect to, for example, outer code rate as in the DVB-NGH standard [15].

Future work consists of generalizing our construction to create well-performing families of rotation matrices $\mathbf{Q}_n(t)$ for all n , not just those which are a power of 2, and additionally improving the performance of our scheme for dimension $n = 8$. The optimal parameter t should also be investigated in connection with outer error-correcting codes. Secondly, we wish to study the decoding complexity of our rotation matrices, and perhaps construct some variants which offer provably simpler sphere decoding complexity. Lastly, we plan on investigating local diversity from an information-theoretic perspective to explain exactly why locally fully diverse constellations can outperform those which are globally fully diverse. We hope that the resulting intuition can be applied to the study of other channels, for example MIMO channels and Rician channels, and not simply the SISO Rayleigh fading channel. We believe the robustness of our approach, especially the general independence of the numerical methods on the particular objective function we used, will prove that our ideas will be similarly useful in other settings.

REFERENCES

- [1] J. Boutros and E. Viterbo, “Signal space diversity: A power- and bandwidth-efficient diversity technique for the Rayleigh fading channel”, *IEEE Trans. on Inf. Theory*, vol. 44, no. 4, July 1998.
- [2] J. Boutros, E. Viterbo, C. Rastello, and J.-C. Belfiore, “Good lattice constellations for both Rayleigh fading and gaussian channels”, *IEEE Trans. on Inf. Theory*, vol. 12, no. 2, March 1996.
- [3] F. Oggier and E. Viterbo, “Algebraic number theory and code design for Rayleigh fading channels”, *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 3, pp. 333–416, 2004.
- [4] J. L. Massey, “Capacity, cutoff rate, and coding for a direct-detection optical channel”, *IEEE Trans. on Comm.*, vol. 29, no. 11, pp. 1615–1621, November 1981.
- [5] G. Forney and G. Ungerboeck, “Modulation and coding for linear gaussian channels”, *IEEE Trans. on Inf. Theory*, vol. 44, no. 6, pp. 2384–2415, October 1998.
- [6] A. Fàbregas and G. Caire, “Multidimensional coded modulation in block-fading channels”, *IEEE Trans. on Inf. Theory*, vol. 54, no. 5, May 2008.

- [7] S. Herath, N.H. Tran, and T. Le-Ngoc, “Rotated Multi-D constellations in Rayleigh fading: Mutual information improvement and pragmatic approach for near-capacity performance in high-rate regions”, *IEEE Trans. on Comm.*, vol. 60, no. 12, December 2012.
- [8] A.O. Hero and T.L. Marzetta, “Cutoff rate and signal design for the quasi-static Rayleigh-fading space-time channel”, *IEEE Trans. on Inf. Theory*, vol. 47, no. 6, September 2001.
- [9] K. Boullé and J.-C. Belfiore, “The cutoff rate of time correlated fading channels”, *IEEE Trans. on Inf. Theory*, vol. 39, no. 2, March 1993.
- [10] A. Edelman, T. Arias, and S. Smith, “The geometry of algorithms with orthogonality constraints”, *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.
- [11] J. Zoellner and N. Loghin, “Optimization of high-order non-uniform QAM constellations”, in *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2013.
- [12] J. Hossain, A. Alvarado, and L. Szczecinski, “BICM transmission using non-uniform QAM constellations: Performance analysis and design”, in *IEEE International Conference on Communications (ICC)*, 2010.
- [13] W. Betts, A. Calderbank, and R. Laroia, “Performance of nonuniform constellations on the gaussian channel”, *IEEE Trans. on Inf. Theory*, vol. 40, no. 5, September 1994.
- [14] L. Zhong, F. Alajaji, and G. Takahara, “Error analysis for nonuniform signaling over Rayleigh fading channels”, *IEEE Trans. on Comm.*, vol. 53, no. 1, January 2005.
- [15] “Next generation broadcasting system to handheld, physical layer specification (DVB-NHG)”, http://www.dvb.org/resources/public/standards/A160_DVB-NHG_Spec.pdf, November 2012.
- [16] M. Lachlan and D. Gomez-Barquero, “Modulation and coding for ATSC 3.0”, in *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2015.
- [17] “Implementation guidelines for a second generation digital terrestrial television broadcasting system (DVB-T2)”, February 2012, <http://www.dvb.org/technology/standards>.
- [18] D. Gómez-Barquero, *Next Generation Mobile Broadcasting*, Taylor & Francis Group, LLC, 2013.
- [19] F. Oggier, J.-C. Belfiore, and E. Viterbo, “Cyclic division algebras: A tool for space-time coding”, *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 1, pp. 1–95, 2007.
- [20] A. Sakzad, A.-L. Trautmann, and E. Viterbo, “Cross-packing lattices for the Rician fading channel”, in *IEEE Information Theory Workshop (ITW)*, 2015.
- [21] A. Sakzad and E. Viterbo, “Full diversity unitary precoded integer-forcing”, *IEEE Trans. on Wireless Comm.*, vol. 14, no. 8, April 2015.
- [22] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, “SCMA codebook design”, in *IEEE Vehicular Technology Conference (VTC)*, 2014.
- [23] D. Karpuk and C. Hollanti, “Rotating non-uniform and high-dimensional constellations using geodesic flow on Lie groups”, in *IEEE International Conference on Communications (ICC)*, 2014.
- [24] D. Karpuk and C. Hollanti, “Multi-dimensional and non-uniform constellation optimization via the special orthogonal group”, in *IEEE Information Theory Workshop (ITW)*, 2014.
- [25] A. Fàbregas, A. Martinez, and G. Caire, “Bit-interleaved coded modulation”, *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 1-2, pp. 1–153, 2008.
- [26] E. Viterbo, “Optimal rotations for number field lattices”, <http://www1.tlc.polito.it/~viterbo/rotations/rotations.html>.
- [27] B. Hall, *Lie groups, Lie Algebras, and Representations: An Elementary Introduction*, Springer, 2003.